The Black Box Society: The Secret Algorithms that Control Money and Information, by Frank Pasquale. Cambridge: Harvard University Press, 2015. 320 pp. ISBN 978–0674368279

Alan Rubel, University of Wisconsin-Madison

There are no shortage of social, ethical, and political questions surrounding the vast, fast, and pervasive flows of information in contemporary society. In *The Black Box Society*, Frank Pasquale makes the case that many of those questions share important features. Specifically, he argues that the algorithms that govern individual reputation, information seeking and retrieval in the finance, insurance, and banking industries are some combination of too powerful, not subtle enough, and too obscure. "Black box," according to Pasquale, is a metaphor that invokes both the data-recording devices used in airplanes and "system[s] whose workings are mysterious" (3). Accordingly, a black box society is one in which people are closely monitored by governments and institutions, and in which people do not comprehend just how much information is collected, how it is used, and what the consequences of information collection and use will be (3). Pasquale's accomplishment in the book is connecting several forms of obscurity (actual secrecy, legal secrecy, and deliberate obfuscation) to broader social phenomena.

For example, in chapter two, Pasquale addresses ways in which information about individuals is collected, aggregated, shared, analyzed, and used in commerce, employment, healthcare, law enforcement, and security. He begins in territory familiar to many: credit bureaus that collect information on individuals from a range of sources and use proprietary algorithms to calculate a credit score, which other entities use as a basis for decisions to extend credit. Pasquale notes this is nothing new, and that today's faceless and automated credit scoring helps displace some of the subjectively biased and discriminatory practices of the past. But the algorithms used to create credit scores are secret, sensitive, and proprietary, and they can change dramatically in ways that individuals cannot fully anticipate or understand. The main concern, however, is that the techniques familiar in credit reporting have extended to other domains, such as marketing and employment, which are not subject to the same regulatory scrutiny as credit reporting.

Pasquale identifies several problems with such pervasive information collection and societal "scoring." The least compelling of these is that it is "creepy." Maybe so, but that is an intuition that requires an argument to provide a reason for reforming or regulating information flows. Information consolidation also exacerbates problems of information insecurity; where lots of well-sorted and useful personal information is held by one party, security breaches are all the more harmful. Such breaches themselves may be "black boxed" insofar as we are unaware of their extent and our vulnerability. Perhaps the most important criticism is people's tendency to trust algorithms, even though they may well be based on prior discrimination, which in turn may reinforce and hide that discrimination.

 Business Ethics Quarterly 26:4 (October 2016). ISSN 1052-150X pp. 568–571
DOI: 10.1017/beq.2016.50

The next target is online searching, which Pasquale argues is susceptible to problems in four areas: "transparency, competition, compensation, and control" (61). The processes used in search include both ranking algorithms and human interventions, for example, in approving apps for sale in the Apple store, but they are all legally protected secrets. This lack of transparency keeps people from knowing why certain apps are rejected from Apple's app store. Are they rejected based on a fair application of standard policy, motivated by politics (in the case of an Apple iPhone app mapping US drone strikes), or an attempt to limit unfavorable publicity (in the case an app calling attention to the spate of suicides by workers at an Apple contractor's factory)?

A related concern is Google's market dominance and its influence over what information people access online. Google keeps the algorithms it uses to rank search results secret in order to prevent people from "gaming" them. But such secrecy "devastates our ability to understand the social world Silicon Valley is creating" (66). That's a problem in that some potentially competing businesses may not break into Google's unpaid search results and yet be unable to tell whether that's because of the algorithm itself, or a manipulation that allows Google to purchase the business at a deflated price. Lack of transparency matters for individuals as well. At times Google will intervene in order to avoid some material from showing up near the top of search results (e.g., Holocaust-denial sites), but at other times fails to intervene (e.g., allowing scurrilous rumors about individuals to remain at the top of results). What accounts for the difference remains opaque. Likewise, with personalized searching: the same search terms used by people at different times or places, or with different search histories yield different results. This may create a "filter bubble" in which people's antecedent views and interests shape search results, which in turn may reinforce those views and interests (79).

The third target is finance, where Pasquale identifies two types of black boxes: "obfuscation in the service of illegality, and opacity resulting from complexity" (103). Pasquale's leading example in this discussion is the use of computer models for mortgage-backed securities in the years before the 2008 financial crisis. Such models had several key problems. First is that they failed to adequately build in risks of such securities. More important for the black box analysis is that they were so complex that no individual analyst (much less an individual investor) could possibly understand the securities and the models, and yet they were subject to "automation bias," or a misplaced confidence in the efficacy of the outputs of algorithms. And that obscurity and bias allowed for wishful thinking about financial markets and "opportunist misuse of models to validate sharp business practices" (107). This combination played an important role in AIG's solvency problems, the LIBOR scandal and others. Pasquale makes a compelling case that it continues to obscure the dubious value of high-frequency trading.

The scope of *The Black Box Society* is impressive. Pasquale connects obscurity across three large domains and explains how that obscurity compounds issues related to credit, advertising, employment, law enforcement, and security (in the case of reputation), to developers' abilities to get their apps in Apple's store, free and paid search results, manipulation of search results, and market dominance (in the case of search), and to fraud and instability in financial markets (in the case of finance).

This is a key contribution in that the obscurity issues that Pasquale describes are important in large part *because* of other facts.

The extent of the terrain covered and the emphasis on the black box metaphor, however, raise a couple of important questions. The first is just what work the black box metaphor is doing. In each of Pasquale's target cases there are important considerations that are not specifically about obscurity. For example, digital reputation in the context of employment is important in part because of weak labor markets, stagnant wages, and unequal bargaining power. Obscurity in search is important in large part because of the market dominance (and potentially anticompetitive behavior) of Google and Apple. And obscurity is important in finance because of its outsized share of (and effect on) the US economy and light regulation. Moreover, as Pasquale explains, the obscurity in each domain has different sources (e.g., trade secrets, complexity, deliberate obfuscation). It is unclear, then, whether black boxes in reputation, search, and finance share important features *other than* obscurity. To put it slightly differently, because the obscurity in each case is important for different reasons and arises from different sources, it is unclear why we should consider them part of the same phenomenon.

That raises a second, closely related question about the *moral salience* of "black-boxiness." One of the key themes of the book is that obscurity functions within larger contexts, and those larger contexts create the conditions under which obscurity matters. As a result, the importance of transparency would seem to differ across cases. Further, different accounts of the moral value of transparency could lead to different prescriptions even within Pasquale's cases. Take reputation. Greater transparency might be justified because people could in some cases exert greater control over their reputation and derive some benefit from doing so, or it might be justified on the grounds that respecting individuals requires that their lives not be governed by secret algorithms. That's an important difference. The first version turns on an empirical claim about what people would do with greater transparency; if good consequences of greater transparency don't pan out, then we would need some other justification for requiring greater transparency. Moreover, it is entirely possible that a more effective means of achieving the desired ends would not involve transparency to individuals at all, but a more thoroughgoing regulatory regime. But the second potential rationale, respecting individuals, would require greater transparency to individuals *regardless* of whether it would actually redound to their material benefit.

The same issue arises in search. One possibility is that greater transparency is important so that individuals can simply understand how their search results are individualized on the grounds that people have an autonomy interest in understanding how important features of the social world work. In that case, transparency policies are justifiable regardless of whether they would alter people's behaviors. Another possibility is that manipulation of search results is problematic for instrumental reasons—for example, because they are anticompetitive or because they have insalubrious effects on democratic life. In that case, transparency may be less important than public intervention of some sort.

This question about the moral salience of black-boxiness is reflected in Pasquale's recommendations. He writes that "[t]he more the black boxes of corporate practices in these areas are revealed, the more pressure will mount to change them" (140). Thus, the principal argument appears to be that greater transparency is instrumentally valuable in rectifying practices that are problematic for reasons other than being obscure. And indeed, many of Pasquale's recommendations are for reforms that are not at root about transparency; for example, he outlines a number of ways to make reputational scoring less arbitrary and more fair by limiting uses of certain kinds of information in hiring and employment decisions or allowing greater reach for antidiscrimination laws. In the case of search, recommendations include allowing a greater role for the Federal Trade Commission in assessing whether app developers or competitors are treated fairly, and a greater potential for legal action for anticompetitive actions. For finance, Pasquale recommends a greater dedication to understanding and policing a range of practices, a greater willingness to prosecute financial crimes, and better funding for agencies to carry out their regulatory work. Of course many of these recommendations might benefit from substantial transparency, at least with respect to regulators. And as Pasquale is careful to explain, transparency alone cannot solve the problems he outlines in reputation, search, and finance. But that leaves us with the concerns expressed above: what work is the black box metaphor doing, in what ways (if any) are the cases similar other than obscurity, and what exactly are the underlying values that such policy prescriptions advance?

In the end, Pasquale makes a compelling case that there are complex and deep problems in reputation, search, and finance, and that those problems are particularly acute because the algorithms governing them are "black boxes." It should also prompt scholars of business ethics and political philosophy to consider some remaining questions about the moral weight of transparency in each of these contexts.